基于 LDA-HMM 的国内创新医疗器械主题演化与技术知识创新管理研究

王雅晴[1]，王琰[1]，杨爽[1]，刘曙霞[1]，胥美美[1]，安新颖[1]
1 中国医学科学院医学信息研究所 北京 100020

摘要：[目的/意义]医疗器械产品是临床科研成果转化和知识创新管理的产物。作为科技产业竞争的主要赛道之一，医疗科技实力是衡量国家总体科研能力和环境的重要方面。知识创新驱动产业发展，产品结构反映知识管理的层次水平，而涉及前沿科技的知识管理与国家战略规划息息相关。从产品主题挖掘与演化分享的角度，透视领域技术知识管理现状，对企业和国家来说具有重要战略意义。[方法/过程]本研究采用了一种综合的技术主题识别和分析方法。通过应用机器学习领域的 LDA 算法，从医疗器械产品数据中识别并聚类潜在的技术主题，进而分析各时期技术主题的分布特征和变化规律。结合隐马尔可夫模型定量预测未来的技术趋势。采用这种组合方法，对 2014-2023 年创新医疗器械的主题分布、演化规律及未来趋势进行了深入的研究。[结果/结论]采用国内获批的创新型医疗器械产品数据识别和整理出 38 个技术主题。结果显示，基因检测和植入电刺激器技术的深入开发。疾病诊断、手术导航和生物补片是未来技术创新和产品竞争的重点。但从数量特征上看，我国研发力量较为薄弱，创新能力处于弱势地位，需引起企业技术研发机构和政府重视。

关键词：竞争情报；知识管理；成果转化；LDA；HMM

分类号：G251

1 引言/Introduction

技术是产品质量和性能的保证，也是产品创新的基础，产品是技术创新的基础，挖掘产品的技术支撑，分析技术主题分布及演变的规律[1]，在一定程度上有助于辅助技术机会预测，对于企业可助力新技术预测，避免研发成本浪费，缩短研发周期[2]，争取竞争优势[3]。对于国家可明确特定领域技术突破方向，尽早铺开前期研究计划，另一方面有助于发现技术缺口，探索新路径，保证国家的战略利益。

医疗器械作为医学技术和医学基础研究的显化产品，从医学学科发展上具有高度的战略意义，一是医疗器械作为医学领域科技成果转化的重要输出产品其对学科发展具有带动性，二是从医学应用上看，医疗器械的推陈出新能够更好地满足临床的需要。从创新的角度来看，诸如医疗器械此类的一众医学学科产品的创新需要驱动着整个医疗行业的产学研合作，成为了医学技术创新的重要组织形式[4]。

国内创新医疗器械从专利、产品定型、原理首创技术先进临床显著价值三个方面去界定，真正的牵引力量是显著的临床价值。在充分理解现有临床诊疗现状如技术、产品的基础上能洞察到临床未被满足的需求，运用企业的技术攻关能力，在医疗器械研开发体系的框架下推陈出新。与美国相比，国内企业对于需求洞察与满足的能力缺乏，甚于技术能力不足阻碍发展。企业模式仍把器械药物化，诉诸基础科研以求解决产品创新问题。基础科研的发展是一切创新的源泉和底层支撑，然而在医疗器械产业界，当前需将临床需求摆在牵引的位置上进一步匹配技术攻关。对当前国内创新医疗器械技术主题状态及演进进行科学分析具有观测价值，本研究以近十年国内获批创新医疗器械为例，探讨创新型医疗器械技术主题、现状及其演化趋势。

专利文献作为技术的主要记录文本之一，因其数据库庞大和较为权威等特点使其能在一定程度上反应技术属性，常作为技术机会分析等技术研究的数据源。通过专利文献挖掘技术主题是技术分析的基础，以往研究常用的方法有 SAO 语义分析[5]、主题模型[6]以及主题聚类[7,8]。以时间维度纵向对技术主题进行趋势分析是研究的主流方向，以往研究使用了灰色预测[9]、技术生命周期[2]、集对分析[10,11]、技术路线图[12]、时间序列分析[13]等方法。然而借助工具的趋势分析往往局限在数据的表层，难以真正联系技术内容，研究结论

与现实之间具有较大偏倚，加之科学和技术的进步本身具有随机性[14]，现有的技术趋势分析和预测研究可能存在着对技术发展过于依赖以往研究延续性的误解。

本研究采用LDA主题模型和HMM隐马尔科夫过程相结合的方法对我国获批的创新医疗器械产品数据进行处理。使用机器分析和标注同时调整模型参数对产品的技术主题进行提炼，通过隐马尔可夫过程对主题词共现矩阵进行建模，获取主题词共现的概率分布，从而得到医疗器械产品技术的主题演化情况，通过热力图对结果进行可视化呈现，解读技术趋势，为医学学科的技术研究和产业进步提供一定的支持。

2 数学模型/Mathematical model

LDA（Latent Dirichlet Allocation）主题模型由 Blei David M.、Ng Andrew Y.和 Jordan[15]于 2003 年提出，用于推测文档的主题分布。该方法可将文本集中每篇文档的主题以概率分布的形式给出，通过分析部分文档抽取其主体分布后，根据主题分布进行聚类或分类，其通过限制主题数量以避免参数过多和过拟合的问题。超参数狄利克雷先验分布如下，dk 代表文本 d 在所有技术主题 k 中的分布：

$$\mathrm{Dir}(\theta_d|\alpha)=\frac{\Gamma\left(\sum_{k=1}^{K}\alpha_k\right)}{\prod_{k=1}^{K}\Gamma(\alpha_k)}\prod_{k=1}^{K}\theta_{dk}^{\alpha_k-1}$$

本文假设使用的医疗器械产品数据服从该分布，在此过程中聚类的主题个数需要研究者自行设置，因此需要计算出该参数的最优选择。Blei[16]于 2012 年提出困惑度指标，通过调试聚类主题数量使得困惑度值最小从而得到最优的主题数量。但这种方法得到的主题数可能造成主题之间差距较小的问题。Teh 等人提出了层次狄利克雷过程方法 (Hierarchical Dirichlet Processes) , 它选择使用一种非参数模型通过自动训练得到最优的主题数[17]。

考虑到模型的泛化能力和主题抽取效果，我们采用了 Perplexity 方法来计算最优的主题数量，同时考虑了困惑度和主题间的相似性。这种方法可以帮助我们更好地确定主题的数量，从而更好地理解和分析文本数据。

在得到主题聚类后，我们通过热力图观察各技术主题随时间的变化趋势。为了进一步得到技术的演化路径，我们采用了一个隐马尔可夫模型得到技术主题的共现概率矩阵来进行深入的技术主题演化分析[18]。流程如下：

（1）设定随机过程中生成的隐状态序列为 $Q=\{q_1,...,q_t\}$，隐藏状态的随机转移序列集合为 $S=\{s_1,...,s_h\}$[18]。

（2）设定状态转移概率分布为 $A=\{a_{ij}\}$，$a_{ij}=P\{q_{t+1}=S_j|q_t=S_i\},1\leqslant i,j\leqslant N$[18]。

（3）当状态为 S_i 时，观测变量的概率分布为：$B=\{b_i(v)\}=\{f\{Q_t=v|q_t=S_i\}\}$，其中 Qt 为第 t 种观测随机变量，观测到序列为 $O=\{o_1,...,o_t\}$[18]。

（4）系统初始状态的概率分布为 $\pi=\{\pi_i,1\leqslant i\leqslant N\}$，其中 π_i 为出现状态 Si 的概率[18]。

使用 Baum-Welch 算法[19]对上述 HMM 模型进行参数估计，得到单一最优状态序列[18]。

3 数据处理与主题演化分析/Data processing and topic evolution analysis

本文使用近年来创新医疗器械审批数据，涉及 2014 年 12 月 10 日至 2023 年 5 月 16 日获批的 202 条医疗器械产品数据，对产品名称进行分词得到单词 1020 余个，例如“支架”、“系统”、“冠状动脉”、“CT”、“血流”、“植入式”、“刺激器”、“导航”、“球囊”等，对分词结果抽检发现分词效果较为理想。

使用 LDA 模型的关键是选择合适的参数。首先，LDA 是一种机器学习算法，其学习效果与迭代次数密切相关[20]。迭代 30 次之后，收敛效果并没有明显区别。在权衡运算速度和科学性后我们设置了 30 次的迭代次数。

在实际操作中我们一共计算了划分为 1 到 200 个主题数时对应的困惑度的值。当变量选择为 45 时，困惑度几乎最小，如图 1 所示。因此，我们将 LDA 模型的主题数参数设定为 45，以达到 LDA 模型较优的学习效果和运行效率。

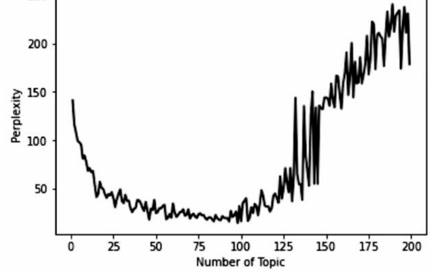


图 1 不同主题数的 Perplexity-Var 值

编码中采用 Python 第三方库对主题信息进行计算，归纳每个主题的关键词和主题名称。排除了与医疗器械技术无关的类别（2 个），去掉由虚词组成的聚类（5 个），最终采用 38 个主题。采用的主题与主题之间避免包容交叉和混淆。具体结果如表 1 所示。

表 1 主题关键词及其命名

No.	KW1	KW2	KW3	KW4	KW5	ID	No.	KW1	KW2	KW3	KW4	KW5	ID
1	基因	突变	探针	测序	试剂盒	基因检测	20	系统	治疗	质子			质子治疗系统
2	消融	冷冻	无菌	导管	低温	冷冻消融设备	21	双源	摄影	系统	体层	数字化	体层摄影设备
3	发生器	脉冲	神经	植入式	电刺激	植入式神经刺激器	22	左心耳	封堵器	系统			左心耳封堵器系统
4	血流	储备	分数	计算	分析	血流储备分数计算	23	医用	直线	加速器	电子	系统	医用电子直线加速器
5	发射	断层	设备	成像	计算机	计算机断层成像设备	24	雷帕霉素	靶向	支架	系统	洗脱	雷帕霉素洗脱支架
6	主动脉	覆膜	输送	支架		覆膜支架	25	盐水	灌注	射频	消融	导管	射频消融导管
7	起搏	经导管	植入式	无导线	系统	植入起搏系统	26	恒温	芯片	扩增	核酸		恒温扩增微控流芯片
8	颗粒	可吸收	氧化	纤维素	止血	氧化纤维素止血颗粒	27	胶原蛋白	软骨	修复	支架		软骨修复支架
9	磁共振	系统	成像	移动式	头颈	磁共振成像	28	患者	程控	充电器			患者程控充电器
10	辅助	检测	监测	软件	诊断	辅助检测/诊断	29	血管内	导管	成像	设备		血管内成像
11	导航	手术	定位	系统	髋关节	手术导航定位	30	角膜	脱细胞	植片			脱细胞角膜植片
12	球囊	扩张	导管	紫杉醇	药物	球囊扩张导管	31	水动力	刀头	治疗	设备		水动力
13	降解	鼻窦	药物	支架	系统	药物支架系统	32	持续	葡萄糖	监测			葡萄糖监测
14	左心室	辅助	植入式	系统		植入左心室辅助	33	钉脚	记忆合金	固定器			记忆合金钉脚固定器
15	晶状体	人工	非球面	衍射		人工晶状体	34	医用	生物	补片	修补		生物补片
16	放射	治疗	调强	伽玛射束	系统	调强放疗治疗计划软件	35	起搏器	临时				临时起搏器
17	多重	静脉	试剂盒	PCR	熔解曲线	多病原体检测试剂盒	36	集成	膜式	氧合器			氧合器
18	超声	诊断	设备	内窥镜		超声诊断仪器	37	取栓	支架				取栓支架
19	瓣膜	主动脉	心脏	介入	系统	人工瓣膜系统	38	周围神经	移植物	修复			周围神经修复移植物

根据实验结果，本文绘制了基于时间变化的医疗器械技术演变趋势热力图（如图 2）。其中，观察图 2 中医疗器械主题的首次出现时间以及各年度变化趋势可知，医疗器械的种类随着科技的进展逐渐多样化。2014-2016 年期间，国内医疗器械关注的创新点仅体现在主题 30（脱细胞角膜植片）、26（恒温扩增微控流芯片）、3（植入式神经刺激器）、1（基因检测），四个研发或应用的方向，通过时间纵向分析，其中主题 3 和主题 1 在十年内均持续保持较高技术产品数量和年度占比份额。主题 9（磁共振成像）、11（手术导航定位）、19（人工瓣膜系统）、30（脱细胞角膜植片）技术获得药品监督局创新医疗器械审批数量较多，在近五年得到充分发展。

脱细胞角膜植片是一种用于眼科手术的人工角膜替代品，它通过去除细胞组织，保留角膜的结构和透明度，用于修复和替代受损的角膜组织，可以治疗如角膜炎等较严重的角膜疾病。它是一种相对较新的技术，首次出现在 20 世纪 90 年代，并逐渐得到了广泛应用。恒温扩增微控流芯片用于基因检测，它的主要原理是将 DNA 样本放在一个小芯片上进行扩增，这种技术的优点是可以在短时间内获得大量的 DNA 样本，并且可以在同一芯片上同时检测多个基因，广泛应用于遗传病筛查、癌症诊断、传染病检测等领域。

植入式神经刺激器用于治疗神经性疾病，它的主要原理是将一种小型设备植入到患者的神经系统中，以缓解疼痛、治疗癫痫等疾病。这种技术可以减轻一些神经性疾病患者的痛苦，提高他们的生活质量。

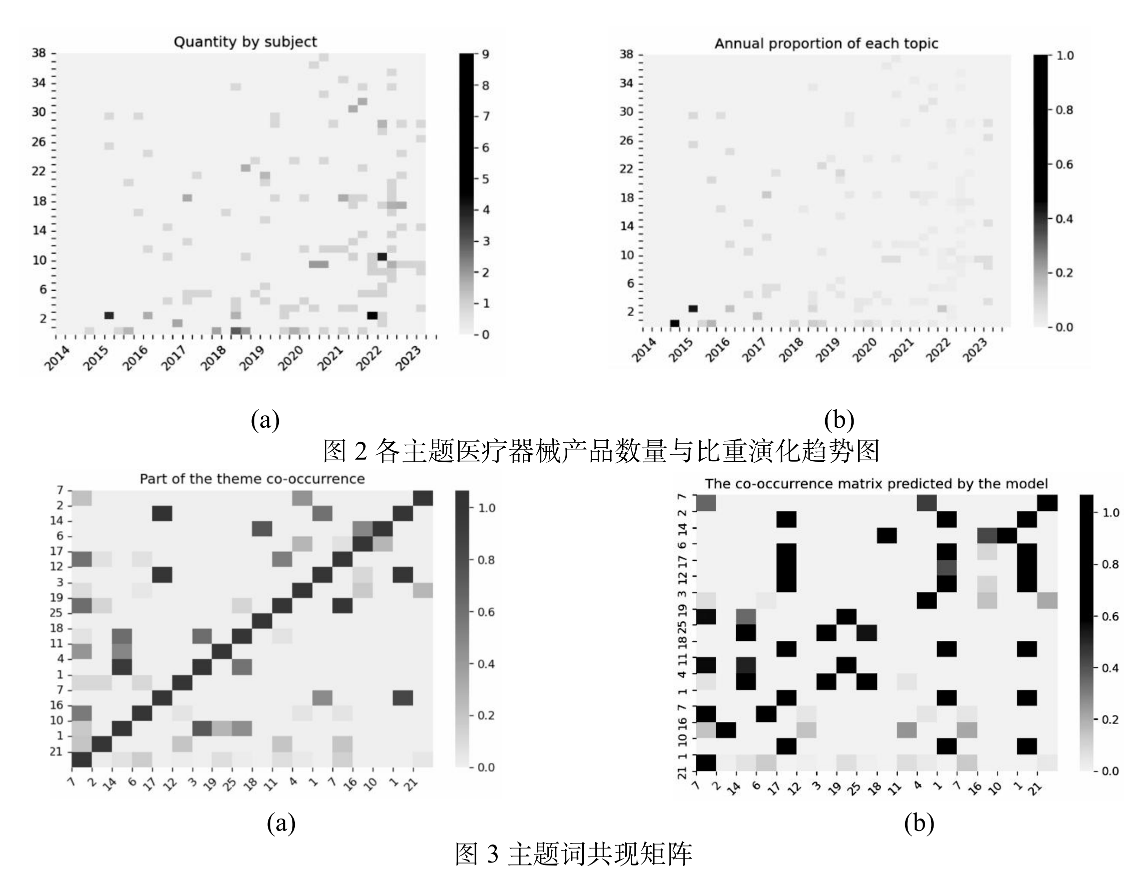
基因检测用于预测个体患病风险，它的主要原理是通过分析个体的基因序列来预测其患病风险。这种技术可以预测一些遗传性疾病的风险，如癌症、心脏病等。它还可以帮助医生制定更加个性化的治疗方案，以提高治疗效果。基因检测技术在近年来得到了广泛应用，特别是在癌症诊断和治疗方面。

磁共振成像是医学影像学中的一种技术，它利用磁场和射频脉冲来生成图像。这种技术的优点是可以生成高分辨率的图像，并且可以用于全身各个部位的检查，可以诊断出许多疾病，如脑部疾病、关节疾病、肿瘤等。

手术导航定位是通过将患者的 CT 或 MRI 图像与手术室中的实时图像进行对比，以帮助医生更加准确地找到病变部位，其可以提高手术的准确性和安全性，减少手术时间提高手术效率。手术导航定位技术在神经外科、

骨科等领域的手术中得到广泛应用。

人工瓣膜系统的主要原理是通过将一种人工瓣膜植入到患者的心脏中，以替代病变的瓣膜。这种技术可以治疗一些严重的心脏瓣膜疾病，如风湿性心脏病、先天性心脏病等。



(a) (b)

图 2 各主题医疗器械产品数量与比重演化趋势图

(a) (b)

图 3 主题词共现矩阵

按照第二部分描述的数学模型，对已经得到的 38 个技术主题的主题词进行筛选，选取 18 个主题词统计其间共现情况，为了更直观地呈现这些相似度，我们构建了一个 18x18 的共现词数对称矩阵，并绘制了热度图（图 3a），由于主题词与自己共现频率最高所以热力图（图 3a）呈现出对角线。将共现矩阵作为初始状态矩阵带入第二部分描述的隐马尔可夫模型中，随后，我们利用 Baum-Welch 算法对 HMM 模型进行了训练，得到了主题词共现概率矩阵（图 3b）。

尽管我们得到了这一结果，但由于技术原因，该结果并非最优解。在未来的研究中，我们计划进一步优化模型参数，以获得更准确的共现概率矩阵。此外，我们还将探讨其他模型和方法，如深度学习模型或聚类分析算法，以更好地处理技术主题的演化过程。

4 创新医疗器械技术知识管理现状/Current situation of technical knowledge management of innovative medical devices

医疗器械产品与医疗技术知识管理之间存在密切的关系。医疗器械作为医疗技术的重要载体，其性能和质量直接影响到医疗技术的应用效果。随着医疗技术的不断创新和发展，医疗器械产品也在不断更新换代，以满足临床需求。同时，医疗器械产品的设计和制造也需要依托先进的医疗技术知识，以确保产品的安全性和有效性。在医疗器械产品的关系理论中，不同产品之间的互补性和竞争性是重要的考量因素。不同类型的医疗器械产品可以相互配合，提供更加全面的医疗服务；而同类产品之间的竞争则有助于推动产品性能的提升和成本的降低。因此，医疗机构在选择医疗器械产品时，需要综合考虑产品的性能、价格、售后服务等因素，以实现最优的资源配置。过去十年我国创新医疗器械技术主题交叉集中在单个诊断或手术技术与智能系统及软件的融合、基因检测、病原检测试剂盒和植入式治疗等技术上。

目前，国内基因检测试剂盒市场已经初具规模，且发展迅速。一方面，国家政策对基因检测试剂盒的发展

给予了大力支持。近年来，政府逐步放开了基因检测试剂盒的生产和销售的限制，推动了基因检测市场的快速发展。另一方面，随着人们健康意识的提高和医疗观念的转变，越来越多的人开始关注基因检测，使得基因检测试剂盒的需求不断增长。国内基因检测试剂盒市场已经形成了以大型医疗设备和试剂生产商为主的竞争格局。其中，华大基因、贝瑞和康、达安基因等公司占据了市场的主要份额。同时，一些新兴的科技公司也在积极布局基因检测领域，如燃石医学、泛生子等。基因检测试剂盒的检测精度和速度将会不断提升。同时，新的基因检测试剂盒将更加注重对特殊疾病的检测，如罕见病、早期肿瘤等。随着精准医学的发展，基因检测试剂盒将会更加注重个体化医疗的需求。个体化基因检测可以帮助患者更好地了解自身健康状况，为个性化治疗提供依据。

我国植入式脑深部电刺激器技术研发在近年来取得了显著的进步。在当前的研发阶段，科研人员正致力于开发更精准、更可靠、更持久稳定的植入式脑深部电刺激器。科研团队在神经工程学、材料科学、微电子技术等领域展开了深入研究，推动了植入式脑深部电刺激器技术的进步。成功研发出具有自主知识产权的植入式脑深部电刺激器，并在精准度、安全性、耐用性等方面取得了重要进展，进一步优化植入式脑深部电刺激器的设计和性能，研究如何提高电刺激器的稳定性和耐用性，以减少故障率和维修率。同时，探索如何更好地适应不同患者的生理需求，提高治疗的个体化程度。如何将先进的神经科学技术如光遗传学、神经调控技术等与植入式脑深部电刺激器相结合也是该主题领域的新课题。上述技术应用将有望进一步提升植入式脑深部电刺激器的治疗效果。在技术趋势方面，我国科研人员正积极探索 5G、物联网等新兴技术在植入式脑深部电刺激器中的应用。通过这些技术的应用，有望实现远程监控、远程调控等功能，提高治疗的便捷性和安全性。此外，人工智能技术的引入也将为植入式脑深部电刺激器的发展带来更多可能性。通过深度学习等技术，科研人员有望开发出能够自动适应患者生理状态、自动调整刺激参数的智能化电刺激器。关注如何将生物材料科学、纳米技术等新兴领域的技术应用于植入式脑深部电刺激器的设计和制造中，有望提高电刺激器的生物相容性、降低免疫排斥反应，同时提高设备的耐用性和安全性。另外，随着精准医疗理念的普及，我国科研人员正致力于开发个体化定制的植入式脑深部电刺激器。通过采集患者的神经影像学数据和生理参数等信息，科研人员将根据每个患者的具体情况定制治疗方案和电刺激器参数。这种个体化定制的方式将进一步提高治疗的精准度和效果。

病原体检测试剂盒是医疗领域中用于快速检测病原体的重要工具，从社会背景角度看，病原体检测试剂盒在公共卫生事件中发挥了至关重要的作用。在传染病疫情爆发时，快速、准确的病原体检测对于防止疫情扩散、保障公众健康具有决定性意义。此外，随着全球人口流动的增加，新发和再次发生的传染病威胁也在不断增加，这进一步凸显了病原体检测试剂盒的重要性。技术研发方面，病原体检测试剂盒经历了不断改进和优化的过程。最早的病原体检测试剂盒主要基于传统的生物学方法，如培养、显微镜观察等。这些方法虽然直观，但耗时较长，且对设备和专业人员的要求较高。随着分子生物学和基因组学的发展，病原体检测试剂盒逐渐转向基于核酸和蛋白质检测的方法。这些方法具有高灵敏度、高特异性和快速等优点，但仍然存在一些挑战，如对痕量样本的检测难度较大、假阳性等问题。近年来，随着生物技术的飞速发展，新型的病原体检测试剂盒不断涌现。其中最具代表性的是基于微阵列芯片和纳米技术的检测试剂盒。这些新型试剂盒将生物识别元件与微电子技术相结合，实现了对多种病原体的快速、并行检测，且具有很高的灵敏度和特异性。此外，基于人工智能和大数据分析技术的智能化病原体检测试剂盒也正在研发中，这些试剂盒将能够自动分析检测结果、提供诊断建议等。未来病原体检测试剂盒将朝着更快速、更准确、更便捷的方向发展。随着技术的不断创新和完善，病原体检测试剂盒的应用领域也将不断拓展。例如，基于便携式设备的移动式病原体检测试剂盒将为偏远地区和紧急情况下的病原体检测提供便利。此外，个性化定制的病原体检测试剂盒将更好地满足特定人群的需求。

2023 年下半年获批的创新医疗器械尚未完全公布，因此，在对未来医疗器械技术主题演化趋势的预测上，本文使用 2023 年前两季度数据作为预测基期。将估测混淆矩阵和转移矩阵参数导入 HMM 模块，得到 2014-2023 年，医疗器械各技术主题演变的隐马尔可夫预测结果。

隐马尔可夫预测结果中，超声诊断技术的比重提升较大，这是由于其无创、便捷、安全、准确等优势。随着医学技术的不断进步，超声诊断已经从传统的二维图像向功能成像发展，能够提供更加全面的医学信息，对于疾病的早期发现、诊断和治疗具有重要意义。此外，随着高端医疗设备的不断涌现，超声诊断技术也在不断创新和完善，能够更好地满足临床需求，为患者提供更优质的医疗服务；基因突变检测仍然是创新医疗器械的主要研发方向，随着人们对基因突变与疾病关系的认识越来越深入，基因突变检测试剂盒的应用范围也越来越广泛，不仅用于疾病的早期诊断和预防，还可用于指导个体化治疗和预后评估。可见，在人口老龄化和慢性病压力下，精准诊断与治疗技术将成为重要的研发方向。

医疗技术的知识管理为医疗器械产品管理提供了丰富的技术支撑和知识资源。通过系统整理、分析和应用

医疗技术知识，医疗机构和研发人员能够更准确地把握医疗器械产品的性能、特点和应用范围，为产品的设计、制造和使用提供科学依据。这有助于确保医疗器械产品的质量和安全，减少医疗事故的发生。医疗器械产品管理则能够推动医疗技术知识的实际应用和转化。医疗器械作为医疗技术的重要载体，其管理涉及产品的注册、备案、质量控制、监督等多个环节。通过严格的产品管理，可以确保医疗器械产品符合相关法规和标准，具备良好的性能和安全性。同时，医疗器械产品的实际应用也为医疗技术知识的完善和创新提供了实践基础和反馈机制。医疗技术的知识管理与医疗器械产品管理的结合还有助于促进医疗行业的多学科合作和资源整合。医疗技术涉及多个学科领域，需要不同专业的知识背景和技能支持。而医疗器械产品管理则需要考虑产品在整个生命周期内的各个环节，涉及多个部门和机构的协作。通过知识管理与产品管理的结合，可以打破学科壁垒，促进不同领域之间的交流和合作，实现资源的优化配置和共享。综上所述，医疗技术的知识管理与医疗器械产品管理相结合的重要性在于它们能够相互支持、相互促进，共同提升医疗服务的效率和质量，保障患者的安全和健康。在未来的医疗行业中，应进一步加强知识管理与产品管理的融合，推动医疗技术的创新和应用，为人类的健康事业作出更大的贡献。

参考文献/References

[1]Kelly K.What technology wants[M].Penguin Books,2010:71-74.
[2]谢志明,张媛,贺正楚,等.新能源汽车产业专利趋势分析[J].中国软科学,2015(9):127-141.
[3]韩震,沈君,曲莎莎.RFID 技术趋势及竞争态势的专利计量分析[J].科研管理,2013(7):11-16.
[4]郭文姣.我国医疗器械产学研医合作创新管理策略研究[D].北京:北京协和医学院,2013.
[5] 李欣,王静静,杨梓,等. 基于 SAO 结构语义分析的新兴技术 识别研究[J].情报杂志, 2016(3): 80-84.
[6] Figuerola C, Marco F, Pinto M. Mapping the evolution of library and information science (1978-2014) using topic modeling on lisa[J]. Scientometrics, 2017,112(3): 1507-1535.
[7] Hu B B, Dong X L, Zhang C W, et al. A lead-lag analysis of the topic evolution patterns for preprints and publications[J]. Journal of the Association for Information Science and Technology, 2015, 66(12): 2643-2656.
[8] Jiang H C, Qiang M S, Lin P. Finding academic concerns of the three gorges project based on a topic modeling approach[J]. Ecological Indicators, 2016, 60: 693-701.
[9] Li W W. Application of grey prediction theory to forecast technology input within the Chinese high-tech industries[C]// Proceedings of the 3rd International Conference on Advanced Computer Control. IEEE, 2011: 88-92.
[10] 李柏洲,李新.基于集对分析的企业技术依赖预警及其演化趋势测度[J].运筹与管理, 2015, 24(2): 262-271.
[11] 黄鲁成,成雨,吴菲菲,等.关于颠覆性技术识别框架的探索[J].科学学研究, 2015, 33(5): 654-664.
[12] 李欣，黄鲁成. 技术路线图方法探索与实践应用研究——基于文献计量和专利分析视角[J].科技进步与对策,2016,33(5): 62-72.
[13] Heo G E, Kang K Y, Song M, et al. Analyzing the field of bioinformatics with the multi-faceted topic modeling technique[J]. BMC Bioinformatics, 2017, 18(Suppl 7): 251.
[14] 官建成.产品创新扩散中的随机现象[J].中国管理科学,1994(3): 44-50.
[15]Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3: 993-1022.
[16]Blei D.Probabilistic topic models[J]. Communications of the ACM, 2012, 55(4): 77-84.
[17] Teh Y W, Jordan M I, Beal M J, et al. Sharing clusters among related groups: hierarchical Dirichlet processes[C]// Proceedings of the Neural Information Processing Systems Conference, 2005: 1385-1392.
[18]陈伟,林超然,李金秋等.基于 LDA-HMM 的专利技术主题演化趋势分析——以船用柴油机技术为例[J].情报学报,2018,37(07):732-741.
[19]Welch L. Hidden Markov models and the Baum-welch algorithm[J]. IEEE Information Theory Society Newsletter, 2003, 53(4): 10-13.
[20]AlSumait L, Barbará D, Domeniconi C. On-line LDA: adaptive topic models for mining text streams with applications to topic detection and tracking[C]// Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. Washington, DC: IEEE Computer Society, 2008: 3-12.

Topic evolution and technology of domestic innovative medical devices based on LDA-HMM Knowledge innovation management

Yaqing Wang[1], Yan Wang[1], Shuang Yang[1], Shuxia Liu[1], Meimei Xu[1], Xinying An[1]
1 Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing, 100020

Abstract： [Purpose/Significance]Medical device products are the products of clinical scientific research results transformation and knowledge innovation management. As one of the main competition tracks in the science and technology industry, the strength of medical science and technology is an important aspect to measure the country's overall scientific research ability and environment. Knowledge innovation drives industrial development, product structure reflects the level of knowledge management, and knowledge management involving frontier science and technology is closely related to national strategic planning. From the perspective of product theme mining and evolution sharing, it is of great strategic significance for enterprises and countries to look into the current situation of domain technical knowledge management.[Method/Process]A comprehensive approach to technical topic identification and analysis is adopted in this study. Through the application of LDA algorithm in the field of machine learning, potential technical topics are identified and clustered from the medical device product data, and then the distribution characteristics and change rules of technical topics in different periods are analyzed. The future technology trend is predicted quantitatively by using hidden Markov model. Using this combination method, the topic distribution, evolution law and future trend of innovative medical devices from 2014 to 2023 are deeply studied.。[Result/Conclusion]38 technical topics were identified and sorted out using domestic approved innovative medical device product data. The results show the deep development of genetic testing and implanted electrical stimulator technology. Disease diagnosis, surgical navigation and biological mesh are the focus of future technological innovation and product competition. However, from the perspective of quantitative characteristics, China's research and development strength is relatively weak, and the innovation ability is in a weak position, which needs to attract the attention of enterprise technology research and development institutions and the government.

Keywords: Competitive intelligence; Knowledge management; Achievement transformation; LDA; HMM